SVOA Paediatrics

e-ISSN 2755-1660

ScienceVolks Open Access

Original Research

Exploring Familial Hypospadias: Genetic Insights from Copy Number Variants in a Quad Family

Sourabh Kumar[1], Jyoti Sharma[2], Dinesh Gupta[3], Rahila Sardar[4], Preeti Lal[4], Vishesh Jain[5], Anjan Kumar Dhua[5], Devendra Kumar Yadav[5], Sandeep Agarwala[5], Monis Bilal Shamsi[6], Hamdi Hameed Almaramhy[7], Harpreet Singh[8], Neeta Kumar[9], Prabudh Goel[5]*

[1]Project Research Scientist I, Department of Paediatric Surgery, All India Institute of Medical Sciences, New Delhi, India.

[2]Project Research Scientist II, Department of Paediatric Surgery, All India Institute of Medical Sciences, New Delhi, India.

[3]Group Leader, Translational Bioinformatics, International Centre for Genetic Engineering and Biotechnology, New Delhi, India.

[4]Vgenomics, India.

[5]Professor of Paediatric Surgery, Department of Paediatric Surgery, All India Institute of Medical Sciences, New Delhi, India.

[6]Centre for Genetics and Inherited Diseases (CGID), Taibah University, Madinah 42361, Saudi Arabia; Department of Basic Medical Sciences, College of Medicine, Taibah University, Madinah 42361, Saudi Arabia.

[7]College of Medicine, Taibah University, Medina, Saudi Arabia.

[8]Scientist F and Programme Officer, Division of Development Research, Indian Council of Medical Research, New Delhi, India.

[9]Scientist, Division of Discovery Research, Indian Council of Medical Research, New Delhi, India.

***Corresponding Author:** Prabudh Goel, Professor of Paediatric Surgery, Department of Paediatric Surgery, All India Institute of Medical Sciences, New Delhi, India, PIN 110029

https://doi.org/10.58624/SVOAPD.2026.05.015

Received: April 24, 2026

Published: June 16, 2026

Citation: Kumar S, Sharma J, Gupta D, Sardar R, Lal P, Jain V, Dhua AK, Yadav DK, Agarwala S, Shamsi MB, Almaramhy HH, Singh H, Kumar N, Goel P. Exploring Familial Hypospadias: Genetic Insights from Copy Number Variants in a Quad Family. *SVOA Paediatrics 2026*, 5:3, 101-114. doi: 10.58624/SVOAPD.2026.05.015

Abstract

Background: The genetic aetiology of hypospadias is likely to be oligogenic with possible interactions between multiple genetic variants and contributory environmental factors. A pathogenic copy number variant (CNV) is usually harboured by 3-14% of patients with rare developmental disorders. With this background, a landscape of CNVs in a family with multiple affected and unaffected progeny is presented with an investigation into the potentially responsible, molecular pathways underlying the etiopathogenesis of hypospadias. The family consists of both parents, two sons with hypospadias, and two unaffected sons (whole exome data unavailable for one unaffected son). CNVkit pipeline was executed and the structural variant files were annotated. The identified CNVs were studied for distribution within the family, inheritance, gene-composition and correlated with available information for potential relevance to the phenotype.

Results: Cumulative analysis (F:father, M:mother, P1-P2:affected progeny, U:unaffected progeny) identified 152 unique CNVs[size:1.49 kb–6.53 Mb) comprising 139 deletions and 13 duplications. P1 and P2 were represented by 29/52 and 22/50 de novo CNVs, respectively. P1 & P2 have 16 common deletion CNVs:8/16 CNVs are absent in U (inherited:6, de novo:2); de novo CNVs: chr6:29100942:29306930:DEL & chr16:11379821:11441076:DEL. de novo CNVs encompass *OR2J1* and *OR14J1* genes expressed in testis and spermatozoa as major histocompatibility complex (MHC)-linked olfactory receptors. CNVs encompassing *GREM1, RRN3, KIAA0753* and *HNF1B* genes relevant to hypospadias were identified.

Conclusion: The landscape of CNVs in familial hypospadias has been presented to enhance the understanding of their distribution, frequency and impact on the development of hypospadias and a database has been generated for future research.

Keywords: *Hypospadias, Whole Exome Sequencing, Copy Number Variants, Single Nucleotide Polymorphisms, OR2J1, OR14J1, GREM1, HNF1B, KIAA0753, RRN3*

Introduction

The etiopathogenesis of hypospadias has been ascribed to genetic composition, endocrine disruptors, maternal factors and aberrations during embryonic development. It is known to have a genetic component with reported heritability of 56.9%[1]. Familial clustering has been described with recurrence in 17% (1 in 6) siblings. Although specific genetic mutations or chromosomal abnormalities have been identified and reported in hypospadias, definitive genes and mutations are largely elusive.

Traditionally, whole exome sequencing (WES) data analysis primarily focuses on single nucleotide polymorphisms (SNPs) WES-based calling, however offers superior sensitivity for pathogenic CNVs[2]. A pathogenic CNV is usually harboured by 3-14% of patients with rare developmental disorders; they might have a crucial role in hypospadias etiopathogenesis, particularly in familial cases or instances wherein traditional single-gene mutations do not fill the gaps in understanding the genetic architecture. Beleza-Meireles et al[3] identified significant enrichment of rare CNVs in patients with hypospadias. Kaminsky et al[4] demonstrated a higher burden of both rare and common CNVs in individuals with developmental disorders, including hypospadias. Chen[5] and van der Zanden[6] have identified CNVs in independent cohorts of hypospadias patients through WES and high-resolution SNP arrays respectively. Familial cases offer a unique opportunity to investigate the inherited CNVs providing insights into the etiopathogenesis and lead to novel pathogenic mechanisms and potential therapeutic targets.

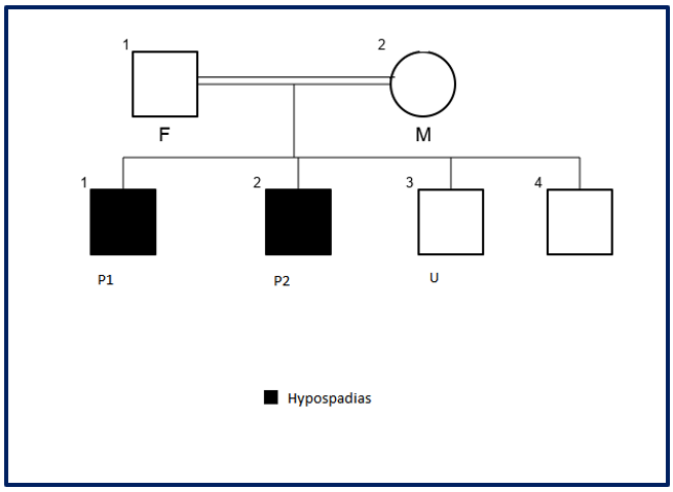
The current synthesis has analysed and presented a landscape of the copy number variants in five family members two of whom are affected with hypospadias, to identify potential molecular etiopathogenetic pathways responsible for the disease.

Material and Methods

The whole-exome sequencing data pertained to five of the six members of a family with hypospadias family (F: father, M: mother, P1 & P2: two siblings with hypospadias respectively, U: unaffected sibling; another unaffected sibling in the family was not subjected to WES) was downloaded from the European Nucleotide Archive [7] (Project: PRJNA982072; SRR24882981-85) [Supplementary Figure 1].

The CNVkit pipeline was executed on the recalibrated BAM files and CNV calling was performed using the CNVkit toolkit [8]. CNVs encompassing genes related to hypospadias were prioritized and visualized using chromosome ideograms. The CNVs were annotated (Annotsv tool) to add ACMG classification, breakpoints annotations, gnomAD, ClinVar, ClinGen, dbVar and regulatory elements annotations.

Literature search was conducted to determine the potential role of short-listed genes in the etiopathogenesis of hypospadias and identify the possible underlying pathways.



Supplementary Figure 1. Family Pedigree [F: Father, M: Mother, P1: Progeny with hypospadias, P2: Progeny with hypospadias, U: Unaffected sibling, 4: another child who was unaffected, sequencing data not available for him).

Results

Landscaping of CNVs in the members of the family: Cumulative analysis across the five samples in the study cohort revealed the presence of 152 unique CNVs (size: 1.49 kB–6.53 Mb, median: 206 kb, IQR: 784.24 kb) [Figure 1] comprising 139 deletions and 13 duplications (Supplementary Table 1).

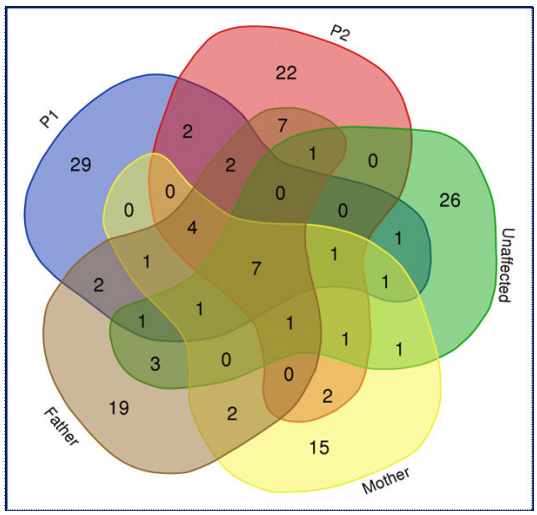
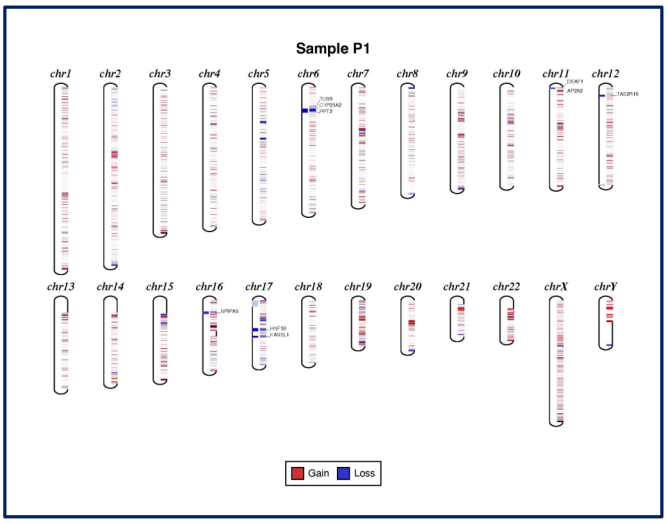


Figure 1. Venn Diagram to elicit the distribution of CNVs in the family.

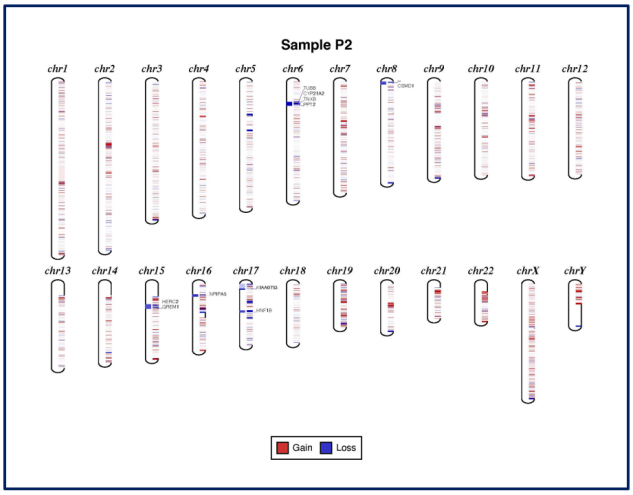
Parents: The father and mother have been represented by 51 (48 deletions & 3 duplications) CNVs and 37 (all deletions) CNVs; 16 CNVs are shared between the two (consanguineous marriage). Seven of these common (16) CNVs have been transferred to all three siblings including two with hypospadias (P1&P2), 4 CNVs have been transferred to affected progeny only (P1 & P2) and 2 CNVs have been transferred to neither. None of the common CNV has been transferred exclusively to the unaffected progeny.

P1 (Affected sibling): Represented by 52 CNVs (49 deletions & 3 duplications). Twenty CNVs have been inherited from one or both parents; 5, 2 and 13 from the father (only), mother (only) and both respectively; all the inherited CNVs were deletions [Supplementary Figure 2a].

P2 (Affected sibling): Represented by 50 CNVs (47 deletions & 3 duplications). Twenty-six CNVs have been inherited from one or both parents; 10 (of which 2 are duplications), 4 (all deletions) & 12 (all deletions) CNVs have been inherited from the father (only), mother (only) and both respectively [Supplementary Figure 2b].



Supplementary Figure 2a. Visualization of CNV distribution across the chromosomes in P1 (phenotype: hypospadias).



Supplementary Figure 2b. Visualization of CNV distribution across the chromosomes in P2 (phenotype: hypospadias)

U (Unaffected sibling): Represented by 45 CNVs (43 deletions & 2 duplications). Eighteen CNVs have been inherited from one or both parents; 5, 4 and 9 have been inherited from the father(only), mother (only) and both respectively; all the inherited CNVs are deletions.

CNVs common to the two affected siblings and not present in the unaffected sibling: P1 & P2 have 16 CNVs in common; all deletions. Of these 8 CNVs are not present in the unaffected sibling and considered highly relevant in context [Table 1].

De-novo CNVs found in the progeny: There were 29 (26 deletions, 3 duplications), 22 (21 deletions, 1 duplication) and 26 (24 deletions, 2 duplications) de novo CNVs (not inherited from the parents) in P1, P2 and U respectively.

The de novo CNVs present in P1-P2 and absent in U included chr6:29100942:29306930:DEL (*OR2J1, OR2J3, OR2J2, LOC124901297, LINC03003, LOC105375006, LOC105375005, OR14J1*) and chr16:11379821:11441076:DEL (genes: *LOC105371082, LOC400499*) [Table 1].

Pathogenicity of the CNVs: Clinical classification (ACMG) identified 21 (of 52) pathogenic CNVs in P1 and 17 (of 50) pathogenic CNVs in P2. Furthermore, 24 and 27 CNVs in P1 and P2 were classified as variants of uncertain significance (VUS)[Supplementary Table 2]. The pathogenic CNVs (n=55; median: 1.37 Mb, IQR: 1.98 Mb) were much larger than benign CNVs (n=16; median: 52.13 kb, IQR: 30.60 kb, $p<0.001$) and those with uncertain significance (n=77, median: 84.08 kb, IQR: 222.01 kb, $p<0.001$) in the study cohort.

Genes involved with the CNVs: The pathogenic CNVs in P1 & P2 encompass a total of 962 and 884 genes respectively of which there are 475 and 423 protein-coding genes respectively[Supplementary Table 2]. The genes involved with the CNVs identified in P1 & P2 and absent in U were identified and enlisted (Table 1). The genes involved with the CNVs shared between the parents and the hypospadiac progeny (F+M+P1+P2) and relevant to the etiopathogenesis of hypospadias were identified and explored for respective function (Table 2).

Chromosomal Localization of pathogenic CNVs: The pathogenic CNVs were preferentially mapped to chromosomes 6 and 17 in both P1 & P2 (42.8% in P1 and 35.3% in P2) (Supplementary Figures 2a&b).

Potential role of CNVs in the etiopathogenesis of Hypospadias: A CNV in P2 viz., chr15:27871862:32969222:DEL (chromosome 15, deletion, q13.1-q13.3) was cross-referenced with a CNV, previously reported in context of 'hypospadias'[9]. This CNV encompasses 95 genes (Supplementary Table 1) of which GREM1 (Gremlin 1, DAN Family BMP Antagonist; 15q13.3) has been previously implicated in the etiopathogenesis of hypospadias [10]. No abnormalities of the GREM1 gene was detected in other family members.

Table 1: List of Copy Number Variants detected in both affected progeny and absent in unaffected sibling.

CNVs	SV:length	CytoBand	Genes
chr5:1295649:1316795:DEL1 *** Shared with both parents**	21146	p15.33	MIR4457
chr5:12410:1294073:DEL1 *** Shared with both parents**	1281663	p15.33	LOC124900928, LOC105374602, PLEKHG4B, LRRC14B, LOC124900929, CCDC127, SDHA, PDCD6-DT, PDCD6, PDCD6-AHRR, AHRR, EXOC3-AS1, EXOC3, LOC124900930, SLC9A3, SLC9A3-OT1, SLC9A3-AS1, LOC107986395, MIR4456, LOC105374606, LOC105374607, LOC105374608, CEP72-DT, CEP72, TPPP, LOC101929898, ZDHHC11B, ZDHHC11, BRD9, TRIP13, LOC124900931, LINC02982, LOC124900932, NKD2, SLC12A7, MIR4635, LOC107986396, TERLR1, LOC124900189, SLC6A19, SLC6A18, TERT
chr6:29100942:29306930:DEL1 *** de novo**	205988	p22.1	OR2J1, OR2J3, OR2J2, LOC124901297, LINC03003, LOC105375006, LOC105375005, OR14J1
chr6:28745393:29100442:DEL1 * Shared with father	355049	p22.1	LOC124901296, LINC01623, HCG14, TRIM27, LINC01556, HCG15, ZNF311, LOC105375002, OR2W1-AS1, OR2W1, OR2B3, OR2J1
chr10:871574:955651:DEL1 *** Shared with both parents**	84077	p15.3	LARP4B, LOC124902363, LARP4B-DT
chr16:14446887:16294299:DEL1 *** Shared with father**	1847412	p13.12-p13.11	PARN, LOC105371094, LOC107984865, BFAR, PLA2G10, NPIPA3, LOC100652777, NPIPA2, ABCC6P2, NOMO1, LOC101927469, MIR3179-1, MIR3670-1, MIR3180-1, PKD1P3-NPIPA1, LOC100288162, MIR6511A1, MIR6770-1, NPIPA1, PDXDC1, MIR1972-1, NTAN1, RRN3, LOC100505915, PKD1P6-NPIPP1, MIR6511B2, MIR3180-4, LOC105371096, LOC105371097, NPIPA5, MPV17L, MPV17L-BMERB1, BMERB1, LOC105371102, MARF1, MIR6506, LOC124903649, MIR484, NDE1, MYH11, LOC124903650, LOC124903651, CEP20, LOC107984869, ABCC1, ABCC6, LOC105371100, NOMO3
chr16:11379821:11441076:DEL1 * de novo	61255	p13.13	LOC105371082, LOC400499
chr17:6636779:6640043:DEL1 *** Shared with both parents**	3264	p13.1	KIAA0753, LOC122526780

Table 2. Genes involved with the CNVs shared between the hypospadiac progeny and their parents considered relevant to the etiopathogenesis of hypospadias.

Development and Morphogenesis	
DEAF1	*DEAF1* encodes a transcription factor important for regulating gene expression during the development of the central nervous system and embryogenesis.
NPIPA5	The *NPIPA5* (Nuclear Pore Complex Interacting Protein Family Member A5) involved in various cellular processes related to the nuclear pore complex (NPC).The NPC is a significant protein structure that controls the movement of molecules between the nucleus and the cytoplasm in eukaryotic cells.
Signal transduction	
CRHR1	*CRHR1* initiates signal transduction upon activation by its ligand, corticotropin-releasing hormone (CRH). This activation triggers a cascade of intracellular events, including the activation of G-proteins and subsequent activation of adenylate cyclase leading to the production of cyclic AMP (cAMP). Elevated cAMP levels then activate protein kinase A (PKA) which phosphorylates various downstream targets, modulating gene expression and cellular responses. Additionally, CRHR1 can activate other signaling pathways such as the mitogen-activated protein kinase (MAPK) pathway which helps the regulation of diverse physiological processes, including stress response, anxiety, and metabolism.
Hormones and Hormonal Regulation	
CYP21A2	*CYP21A2* encodes 21-hydroxylase involved in the regulation of hormones cortisol and aldosterone which are found in the adrenal glands. This enzyme belongs to the cytochrome P450 family involved in various functions such as drug metabolism, cholesterol production, hormone synthesis, and lipid biosynthesis
CRHR1	*CRHR1* gene encodes a G-protein coupled receptor that binds neuropeptides of the corticotropin releasing hormone family that are major regulators of the hypothalamic-pituitary-adrenal pathway. Dysregulation of CRHR1 has been implicated in various stress-related disorders making it a potential therapeutic target.
Gene Transcription and DNA repair	
DEAF1	The *DEAF1*(deformed epidermal autoregulatory factor-1) is a important transcription factor involved in the development of the central nervous system and early embryogenesis. It is predominantly expressed in the brain and central nervous system, where it regulates the expression of numerous genes act as a both activator and a repressor.
MIER2	*MIER2* encodes a protein involved in regulating gene expression. This protein plays a role in chromatin remodeling and transcriptional repression. It is implicated in various cellular processes, including development and differentiation.
Potein Expression, Modification, Transport and Degradation	
AP2A2	*AP2A2* encodes a subunit of the adaptor protein complex 2 (AP-2), which is crucial for clathrin-mediated endocytosis. This process involves the internalization of nutrients and receptors into the cell. AP2A2 plays a significant role in vesicle formation and trafficking within the cells. Proper functioning of AP2A2 is essential for maintaining cellular homeostasis and efficient intracellular transport.
HERC2 * Shared with both parents	*HERC2* gene involved in protein degradation processes within cells. It encodes an E3 ubiquitin ligase enzyme that tags the proteins with ubiquitin molecules marking them for degradation by the proteasome. HERC2 is particularly known for its role in DNA repair and regulation of cell cycle progression. Mutations in HERC2 have been associated with certain genetic disorders and conditions, highlighting its importance in maintaining cellular integrity and function.3264

Table 2 continued...

LMF1	*LMF1* is responsible for encoding Lipase Maturation Factor 1 essential for the proper maturation and function of lipoprotein lipase involved in lipid metabolism.
SPPL2C	*SPPL2C* encodes a member of the signal peptide peptidase-like protease family involved in intramembrane proteolysis and regulation of immune responses.
Membrane proteins, receptors, transporters, and ion channels	
ADAM2	*ADAM2* encodes a member of the ADAM (a disintegrin and metalloproteinase) family of proteins involved in sperm-egg binding during fertilization.
MUC2	*MUC2* responsible for encoding the mucin 2 protein which forms the structural component of mucus secretions, provide the protection and lubrication to various epithelial surfaces in the body mainly in gastrointestinal tract.
MUC5B	*MUC5B* is a major component of respiratory mucus secretions responsible for maintaining airway hydration and clearance of foreign particles and pathogens.
TAS2R19	*TAS2R19* encodes a bitter taste receptor protein that belongs to the TAS2R family of G protein-coupled receptors which are responsible for detecting and responding to bitter compounds in the oral cavity.
Carcinogenesis/Tumor suppression	
CSMD1	*CSMD1* involved in various cellular processes, including cell adhesion, signaling, and immune regulation with implications in cancer development and neurodevelopmental disorders.
USP6	*USP6* encodes a deubiquitinating enzyme involved in regulating protein stability and degradation with potential implications in various cellular processes, including cancer progression and neurological disorders.
Miscellaneous Genes	
GGT6	*GGT6* encodes the gamma-glutamyltransferase 6 enzyme involved in glutathione and xenobiotic metabolism and potentially affect the oxidative stress response and detoxification mechanisms.
KRT33B	*KRT33B* involved in the structural integrity of epithelial cells mainly in the skin and hair.
KRTAP4-9	*KRTAP4-9* potentially involved in hair structure and development.
PPT2	*PPT2* (Palmitoyl-Protein Thioesterase 2) encodes an enzyme responsible for removing palmitate residues from modified proteins. Mutations in this gene can lead to neuronal ceroid lipofuscinosis is a rare inherited lysosomal storage disorder affecting the nervous system.
Uncharacterized Genes	
REXO1	*REXO1* codes for the RNA exonuclease 1 enzyme involved in RNA degradation processes and RNA quality control mechanisms within cells.
TNXB	*TNXB* encodes the tenascin-X protein a component of the extracellular matrix that contributes to the structural organization and integrity of connective tissues throughout the body. Mutations in TNXB are associated with various connective tissue disorders, such as Ehlers-Danlos syndrome.

Table 3. Genes involved with the CNVs chr17:35764425:38850214:DEL and chr17:36454631:37925499:DEL overlap with genomic regions implicated in congenital anomalies of the kidney and urinary tract .

CNVs	Genes	Phenotypes
chr17:35764425:38850214:DEL	MMP28, C17orf50, TAF15, HEATR9, LOC105371745, CCL5, LRRC37A8P, RDM1, LYZL6, CCL16, CCL14, CCL15-CCL14, CCL15, LOC107985068, CCL23, LOC105371746, CCL18, CCL3-AS1, CCL3, CCL4, LOC101927369, LOC107985055, TBC1D3B, LOC128966684, LOC128966706, CCL3L3, CCL4L2, LOC105371747, TBC1D3I, LOC102724956, LOC101060212, TBC1D3G, LOC102723414, TBC1D3H, TBC1D3F, LOC105371749, ZNHIT3, MYO19, PIGW, GGNBP2, DHRS11, LOC107985031, MRM1, LOC105371750, LOC105371751, LHX1-DT, LHX1, AATF, MIR2909, LOC105371753, ACACA, SNORA90, C17orf78, TADA2A, DUSP14, SYNRG, DDX52, MIR378J, LOC105371756, HNF1B, LOC124903989, LOC105371754, LOC105371757, YWHAEP7, TBC1D3K, LOC102723608, TBC1D3L, TBC1D3D, LOC102723819, TBC1D3C, LOC101929950, TBC1D3E, LOC102723851, LOC102723933, TBC1D3, NPEPPSP1, LOC105371760, MRPL45, GPR179, LOC124903991, SOCS7, ARHGAP23, LOC101929494, SRCIN1, LOC105371761, LOC124903992, EPOP, LOC105371763, MIR4734, MLLT6, LOC124903993, LOC105371762, MIR4726, CISD3, PCGF2, LOC100287808, PSMB3, PIP4K2B, CWC25, MIR4727, SPMAP1, LOC124903994, RPL23	17q12 recurrent (RCAD syndrome) region (includes HNF1B);Acetyl-CoA carboxylase deficiency, 613933 (3) AR;Glycosylphosphatidylinositol biosynthesis defect 11, 616025 (3) AR;Night blindness, congenital stationary (complete), 1E, AR, 614565 (3) AR;PEHO syndrome, 260565 (3) AR;Renal cysts and diabetes syndrome, 137920 (3) AD;Type 2 diabetes mellitus, 125853 (3) AD;Renal cell carcinoma, 144700 (3); Renal:cysts:and:diabetes:syndrome; Turnpenny-Fry syndrome, 618371 (3) AD
chr17:36454631:37925499:DEL	LOC105371749, ZNHIT3, MYO19, PIGW, GGNBP2, DHRS11, LOC107985031, MRM1, LOC105371750, LOC105371751, LHX1-DT, LHX1, AATF, MIR2909, LOC105371753, ACACA, SNORA90, C17orf78, TADA2A, DUSP14, SYNRG, DDX52, MIR378J, LOC105371756, HNF1B, LOC124903989, LOC105371754, LOC105371757, YWHAEP7, TBC1D3K	17q12 recurrent (RCAD syndrome) region (includes HNF1B);Acetyl-CoA carboxylase deficiency, 613933 (3) AR;Glycosylphosphatidylinositol biosynthesis defect 11, 616025 (3) AR;Night blindness, congenital stationary (complete), 1E, AR, 614565 (3) AR;PEHO syndrome, 260565 (3) AR;Renal cysts and diabetes syndrome, 137920 (3) AD;Type 2 diabetes mellitus, 125853 (3) AD;Renal cell carcinoma, 144700 (3); Renal:cysts:and:diabetes:syndrome; Turnpenny-Fry syndrome, 618371 (3) AD

Further analysis identified 10 and 9 CNVs in P1 and P2 respectively encompassing genes implicated in hypospadias [11-14] (Supplementary Table 3). Within the eight CNVs common to P1 & P2 and not present in the unaffected sibling (U), two specific loci of deletions, viz., *chr16:14446887:16294299:DEL* (present in F, P1, P2; size: 1.8 Mb; chromosome 16, p13.13) and *chr17:6636779:6640043:DEL* (present in F, M, P1, P2; size: 3.3kb; chromosome 17, p13.1) relating to hypospadias genes *RRN3* (RNA Polymerase I Transcription Factor) and *KIAA0753* respectively were identified. While a single copy of *RRN3* was deleted as part of CNV in P1, the other copy was mutated at position 21800714 (GRCh38) [NC_000016.10:21800713:A:T; rs556085497; A>T; functional consequence: non-coding transcript variant]. The deletion of *KIA0753* was observed across both the alleles.

Annotations revealed that CNV deletions chr17:35764425:38850214:DEL (size: 3 Mb, chromosome 17, q12)[P1, M and U] and chr17:36454631:37925499:DEL (size: 1.47 Mb, chromosome 17, q12)[P2] overlap with genomic regions implicated in congenital anomalies of the kidney and urinary tract (CAKUT). The genes encompassed by these CNVs (Table 3) include those associated with loss-of-function implications in various phenotypes, including developmental anomalies of the kidneys and the urinary tract, renal cell carcinoma, PEHO syndrome[11] and Turnpenny-Fry syndrome [11]. Included amongst these is the HNF1B (Hepatocyte Nuclear Factor 1B/ HNF1 Homeobox B) gene, which is involved in hypospadias [15], CAKUT and RCAD (OMIM 137920) syndrome [16].

Discussion

Almaramhy et al. [17] have reported a novel, missense mutation involving the HSD3B2 gene (Chr1:119964631T>A, c.507T>A, p. N169K) in all six members of the index family; the SNP was homozygous in the affected members while the parents and the unaffected siblings were heterozygous carriers. Consanguineous marriages offer unique opportunities for genetic studies due to the concentration of genetic variations within the families, particularly those with recessive inheritance. The higher frequency of encountering rare alleles in a homozygous state facilitates the identification of mutations underlying the disease.

The current study identified 152 unique CNVs in a family with two affected siblings, highlighting the potential role of these variants in the development of hypospadias. Both inherited and de novo CNVs have been identified, some potentially relevant to hypospadias in a familial setting. Inherited CNVs are usually more prevalent than the de novo ones and contribute to genetic diversity within families and populations. While the deleterious CNVs are eliminated through natural selection, those that are compatible with life may be concentrated in specific populations due to genetic drifts, population bottlenecks and founder effects, historical migrations and geographical isolation. It is usually the de novo CNVs that play a significant role in developmental or genetic malformations. De novo CNVs may arise during gametogenesis or early embryonic development and are present in most of the cells of the organism. Post-zygotic mutations affecting the germ cells in the parental gonad (gonadal mosaicism) may lead to genetically distinct populations of cells in the gonad of an individual[18]. In such situations, a proportion of germ cells may harbour the damaging CNV which is not necessarily transmitted to all progeny. The same family may have siblings with and without the CNV or more than one sibling sharing the so-called de novo CNVs.

The index family presented with a myriad of CNVs including deletions and duplications. Although both types may be associated with developmental delays, intellectual disabilities and congenital anomalies, the deletion CNVs are usually associated with milder traits to severe developmental disorders depending upon their size and location. Contrarily, duplication CNVs present with altered gene dosages leading to gene over-expression, disrupted gene regulation or novel fusion gene formation.

The presence of three siblings (with WES data) in a family of which two have the hypospadiac phenotype offered a unique opportunity to look into the CNVs potentially relevant to the etiopathogenesis of hypospadias. There were 16 CNVs common to both the affected siblings (P1 and P2) of which 8 were shared with the unaffected sibling. The other CNVs may also be relevant in view of oligogenic or polygenic origin of hypospadias.

Two de novo CNVs, viz., chr6:29100942:29306930:DEL and chr16:11379821:11441076:DEL have been identified in both P1 & P2 suggesting that these new genetic changes may play a crucial role in the etiology of the disease. These CNVs encompass the genes *OR2J1, OR2J2, OR2J3, OR14J1, LOC400499, LOC105371082, LOC105375005, LOC105375006, LOC124901297* and LINC03003. OR2J1 and OR14J1 belong to the large olfactory receptor gene family; they are expressed in the testis and spermatozoa as major histocompatibility complex-linked olfactory receptors [19]. Their presence may be indicative of a potential novel mechanism in the development of hypospadias including abnormalities in olfactory receptor expression or signalling pathways during embryonic development. *OR2J1, OR14J1* and additional 91 olfactory genes together amount to the specific chemoreceptor repertoire contributing to chemotaxis and chemokinesis; their presence in hypospadias suggests additional functions in reproductive tissues. *OR2J1* and *LINC03003* genes have been reported previously in context of sexual dimorphism [20], while *LINC03003*, *OR2J2*, *OR2J3* and *OR14J1* have been reported in context of inguinal hernia [21] and *LOC105375005* in context of endometrial carcinoma [22].

The pathogenic CNVs showed a preferential distribution on chromosomes 6 and 17. Among the genes encompassed by these CNVs, *HNF1B* on chromosome 17 has previously been implicated in hypospadias and other developmental abnormalities of the genitourinary tract. Variants involving chromosome 17 may therefore be of particular interest in understanding the genetic basis of the condition. The CNVs classified as variants of uncertain significance (VUS) identified in the present analysis require further investigation to clarify their biological and clinical relevance. Determining the contribution of these variants to disease development remains important for improving risk assessment, genetic counselling, and clinical management of affected individuals and their families.

The analysis identified that the pathogenic CNVs and those with uncertain significance were much larger than the benign CNVs. The finding is congruent with standard knowledge since larger CNVs are likely to involve a larger number of genes including the regulatory regions, adversely affecting critical cellular functions. Simultaneous loss of multiple genes affecting common or related pathways is further relevant in oligogenic and polygenic disorders like hypospadias. The smaller CNVs are not necessarily benign; those affecting single genes may be highly pathogenic too if the gene is indispensable for normal development and function.

The protein-coding *GREM1* gene is a part of the DAN family of the bone morphogenetic protein (BMP) antagonist family. It is known to play a role in cell differentiation, proliferation and apoptosis during embryonic development, affecting organogenesis, body patterning and tissue differentiation. The functional implications of this gene relate to the renal and central nervous system development [23]. The gremlin protein relays the sonic hedgehog (SHH) signal from the polarizing region to the apical ectodermal ridge [24]. *SHH* is involved in the growth and differentiation of the genital tubercle and may be relevant for the development of external genitalia including the penis. *GREM1* may also affect the development of urethra and external genitalia through disruption in BMP signalling. Polymorphisms in the GREM1 gene have been linked to increased risk of hypospadias in European and Southern Han Chinese populations [10]. It has also been implicated in the formation and development of the other parts of the urogenital system.

Two CNVs, viz., *chr16:14446887:16294299:DEL* and *chr17:6636779:6640043:DEL* directly disrupting the *RRN3* and *KIAA0753* genes respectively were identified in the affected individuals but absent in the unaffected sibling, strengthening their association with the disease. The *RRN3* (RNA Polymerase I Transcription Factor/ Transcription Initiation Factor IA/ TIF-IA) is a protein-coding gene that plays a crucial role in ribosomal RNA transcription [25]; its disruption has been implicated in sexual dimorphism [20], Treacher Collins Syndrome I with under-developed external genitalia[26] and Diamond-Blackfan anemia[26]. The presence of CNVs in *RRN3* in the affected siblings but not in the unaffected sibling strengthens its potential role in hypospadias. Renal abnormalities and hypospadias have been reported in individuals affected with Diamond-Blackfan anaemia [27]. The protein-coding *KIAA0753* gene is a component of the origin recognition complex (ORC), essential for DNA replication initiation and cell cycle regulation processes critical to genitourinary system formation. The encoded protein regulates ciliogenesis, cilia maintenance and centripolar duplication. Mutations in this gene have been associated with skeletal ciliopathies, autosomal recessive orofaciodigital syndrome XV [60] and autosomal recessive Joubert syndrome 38[28]. The gene has also been associated with hypospadias [29-30] and sexual dimorphism [20].

The HNF1B gene encodes the transcription factor 1B which is relevant to the embryological development of several tissues including the kidneys. Congenital anomalies of the kidney and urinary tract including bilateral cystic dysplasia, structural anomalies of the genitalia, and cysts of the seminal vesicles or the epididymis have been described in affected individuals.

Both the parents and the unaffected sibling did not manifest the hypospadias phenotype despite potentially carrying some of the identified genetic variants (*KIAA0753, GGT6, USP6, DPH1* in chr17:108342:6635103:DEL and *KIAA0753* in chr17:6636779:6640043:DEL). This observation supports the oligogenic and polygenic theory of inheritance; multiple genetic variants interact in a complex manner and contribute to the disease phenotype, role of incomplete penetrance, the presence of genetic modifiers that modulate disease manifestation. This notion is further reinforced by the observation that the pathogenic CNVs were significantly larger than the benign CNVs, indicating that larger CNVs affecting multiple genes are more likely to be deleterious. The responsible genetic variants present in the affected individuals may work in unison towards hypospadias development through complete disruption of the underlying molecular pathways responsible. Identification of CNVs of uncertain significance (VUS) highlight the challenges in interpreting the clinical significance of CNVs and the need for comprehensive functional studies to elucidate their roles in disease development.

While the analysis is limited to datasets pertaining to a single family, it has successfully created a comprehensive database of CNVs and respective genes potentially relevant to the etiopathogenesis of the condition. Studying CNVs in a single family can uncover rare variants that may not be easily detectable in larger population studies and provide unique insights. The availability of a detailed phenotype, particularly the anatomic differences between the two involved brothers, shall assist in correlating the phenotypic differences with the differentially expressed CNVs between the two affected brothers. The study can serve as a starting point to uncover similar CNV patterns in other affected families and validate these findings through their impact on gene expression or protein function thereby unlocking crucial insights and driving our understanding on genetic basis of hypospadias further. Last but not the least, while CNVs and SNPs may contribute to the genetic architecture of congenital malformations, a comprehensive understanding of either is an indispensable step towards genetic counselling, risk assessment and personalized medicine or potentially targeted therapies.

Conclusions

The present study has presented a landscape of CNVs in hypospadias in an effort to understand their distribution, frequency and impact on the development of hypospadias and generated a database for future research. The identification of novel CNVs and their potential pathogenic mechanisms provides valuable insights into the genetic underpinnings of this condition and paves the way for improved diagnostic and therapeutic strategies. The index family was optimal for such analysis in view of consanguineous marriage, availability of WES data in a quad family with more than one affected progeny and unaffected progeny as additional control, thereby enabling identification of de novo and inherited CNVs along with genotype-phenotype correlation.

The study has identified 2 de novo CNVs affecting the OR2J1 and OR14J1 genes expressed in testis and spermatozoa as major histocompatibility complex (MHC)-linked olfactory receptors. Additional CNVs encompassing genes known to be relevant to hypospadias such as GREM1, RRN3, KIAA0753 and HNF1B have also been identified and landscaped. Genes involved with the etiopathogenesis of hypospadias have been identified in the parents (phenotype: normal) supporting the theory of oligogenic or polygenic inheritance. Further research is needed to validate these findings and explore the functional implications of the identified CNVs in the pathogenesis of hypospadias.

Acknowledgement

We thank the family investigated for their invaluable contribution to literature.

Ethics Statement

Public domain data downloaded from European Nucleotide Archive has been re-analysed hence ethical considerations are not relevant herein.

No photographs requiring consent for publication have been included.

All authors have contributed significantly to the manuscript as per the ICMJE guidelines.

Conflict of Interest

None of the authors have any conflict of interest

Funding

The study has not received any funding

References

1. Stoll C, Alembik Y, Roth MP, Dott B. Genetic and environmental factors in hypospadias. J Med Genet. 1990 Sep;27(9):559-63. doi: 10.1136/jmg.27.9.559.

2. Danecek P, Gardner EJ, Fitzgerald TW, Gallone G, Kaplanis J, Eberhardt RY, Wright CF, Firth HV, Hurles ME. Detection and characterisation of copy number variants from exome sequencing in the DDD study. Genetics in Medicine Open. 2024 Jan 28:101818.

3. Beleza-Meireles A, Lundberg F, Lagerstedt K, Zhou X, Omrani D, Frisén L, Nordenskjöld A. FGFR2, FGF8, FGF10 and BMP7 as candidate genes for hypospadias. Eur J Hum Genet. 2007 Apr;15(4):405-10. doi: 10.1038/sj.ejhg.5201777. Epub 2007 Jan 31.

4. Kaminsky EB, Kaul V, Paschall J, Church DM, Bunke B, Kunig D, Moreno-De-Luca D, Moreno-De-Luca A, Mulle JG, Warren ST, Richard G, Compton JG, Fuller AE, Gliem TJ, Huang S, Collinson MN, Beal SJ, Ackley T, Pickering DL, Golden DM, Aston E, Whitby H, Shetty S, Rossi MR, Rudd MK, South ST, Brothman AR, Sanger WG, Iyer RK, Crolla JA, Thorland EC, Aradhya S, Ledbetter DH, Martin CL. An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. Genet Med. 2011 Sep;13(9):777-84. doi: 10.1097/GIM.0b013e31822c79f9.

5. Chen Z, Lei Y, Finnell RH, Ding Y, Su Z, Wang Y, Xie H, Chen F. Whole-exome sequencing study of hypospadias. iScience. 2023 Apr 12;26(5):106663. doi: 10.1016/j.isci.2023.106663.

6. van der Zanden LF, van Rooij IA, Feitz WF, Knight J, Donders AR, Renkema KY, Bongers EM, Vermeulen SH, Kiemeney LA, Veltman JA, Arias-Vásquez A, Zhang X, Markljung E, Qiao L, Baskin LS, Nordenskjöld A, Roeleveld N, Franke B, Knoers NV. Common variants in DGKK are strongly associated with risk of hypospadias. Nat Genet. 2011 Jan;43(1):48-50. doi: 10.1038/ng.721. Epub 2010 Nov 28. Erratum in: Nat Genet. 2011 Mar;43(3):277.

7. Kanz C, Aldebert P, Althorpe N, Baker W, Baldwin A, Bates K, Browne P, van den Broek A, Castro M, Cochrane G, Duggan K, Eberhardt R, Faruque N, Gamble J, Diez FG, Harte N, Kulikova T, Lin Q, Lombard V, Lopez R, Mancuso R, McHale M, Nardone F, Silventoinen V, Sobhany S, Stoehr P, Tuli MA, Tzouvara K, Vaughan R, Wu D, Zhu W, Apweiler R. The EMBL Nucleotide Sequence Database. Nucleic Acids Res. 2005 Jan 1;33(Database issue):D29-33. doi: 10.1093/nar/gki098.

8. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLoS Comput Biol. 2016 Apr 21;12(4):e1004873. doi: 10.1371/journal.pcbi.1004873.

9. Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, Van Vooren S, Moreau Y, Pettett RM, Carter NP. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet. 2009 Apr;84(4):524-33. doi: 10.1016/j.ajhg.2009.03.010.

10. Deng F, Zhao J, Jia W, Fu K, Zuo X, Huang L, Wang N, Xia H, Zhang Y, Fu W, Liu G. Increased hypospadias risk by GREM1 rs3743104 [G] in the southern Han Chinese population. Aging (Albany NY). 2021 May 5;13(10):13898.

11. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res. 2005 Jan 1;33(Database issue):D514-7. doi: 10.1093/nar/gki033.

12. Abnormal Genitalia/ Disorders of Sex Development Panel. Blueprint Genetics Available Online: https://blueprintgenetics.com/tests/panels/endocrinology/abnormal-genitalia-disorders-of-sex-development-panel/ (Accessed on 28 May 2024)

13. Disorders of Sex Development NGS Panel. Fulgent genetics. Available Online: https://www.fulgentgenetics.com/Disorders-of-Sex-Development (Accessed on 28 May 2024)

14. Chang J, Wang S, Zheng Z. Etiology of Hypospadias: A Comparative Review of Genetic Factors and Developmental Processes Between Human and Animal Models. Res Rep Urol. 2020 Dec 24;12:673-686. doi: 10.2147/RRU.S276141.

15. Gambella A, Kalantari S, Cadamuro M, Quaglia M, Delvecchio M, Fabris L, Pinon M. The Landscape of HNF1B Deficiency: A Syndrome Not Yet Fully Explored. Cells. 2023 Jan 13;12(2):307. doi: 10.3390/cells12020307.

16. Devuyst O, Olinger E, Weber S, Eckardt KU, Kmoch S, Rampoldi L, Bleyer AJ. Autosomal dominant tubulointerstitial kidney disease. Nat Rev Dis Primers. 2019 Sep 5;5(1):60. doi: 10.1038/s41572-019-0109-9.

17. Almaramhy HH, Abdul Samad F, Al-Harbi G, Zaytuni D, Imam SN, Masoodi T, Shamsi MB. Identification of a novel candidate HSD3B2 gene variant for familial hypospadias by whole-exome sequencing. Front Genet. 2023 Jun 13;14:1106933. doi: 10.3389/fgene.2023.1106933.

18. Thorpe J, Osei-Owusu IA, Avigdor BE, Tupler R, Pevsner J. Mosaicism in Human Health and Disease. Annu Rev Genet. 2020 Nov 23;54:487-510. doi: 10.1146/annurev-genet-041720-093403.

19. Flegel C, Vogel F, Hofreuter A, Schreiner BS, Osthold S, Veitinger S, Becker C, Brockmeyer NH, Muschol M, Wennemuth G, Altmüller J, Hatt H, Gisselmann G. Characterization of the Olfactory Receptors Expressed in Human Spermatozoa. Front Mol Biosci. 2016 Jan 7;2:73. doi: 10.3389/fmolb.2015.00073.

20. Vosberg DE, Pausova Z, Paus T. The genetics of a "femaleness/maleness" score in cardiometabolic traits in the UK biobank. Sci Rep. 2023 Jun 5;13(1):9109. doi: 10.1038/s41598-023-36132-1.

21. Choquet H, Li W, Yin J, Bradley R, Hoffmann TJ, Nandakumar P; 23 and Me Research Team; Mostaedi R, Tian C, Ahituv N, Jorgenson E. Ancestry- and sex-specific effects underlying inguinal hernia susceptibility identified in a multiethnic genome-wide association study meta-analysis. Hum Mol Genet. 2022 Jul 7;31(13):2279-2293. doi: 10.1093/hmg/ddac003.

22. O'Mara TA, Glubb DM, Amant F, Annibali D, Ashton K, Attia J, Auer PL, Beckmann MW, Black A, Bolla MK, Brauch H, Brenner H, Brinton L, Buchanan DD, Burwinkel B, Chang-Claude J, Chanock SJ, Chen C, Chen MM, Cheng THT, Clarke CL, Clendenning M, Cook LS, Couch FJ, Cox A, Crous-Bous M, Czene K, Day F, Dennis J, Depreeuw J, Doherty JA, Dörk T, Dowdy SC, Dürst M, Ekici AB, Fasching PA, Fridley BL, Friedenreich CM, Fritschi L, Fung J, García-Closas M, Gaudet MM, Giles GG, Goode EL, Gorman M, Haiman CA, Hall P, Hankison SE, Healey CS, Hein A, Hillemanns P, Hodgson S, Hoivik EA, Holliday EG, Hopper JL, Hunter DJ, Jones A, Krakstad C, Kristensen VN, Lambrechts D, Marchand LL, Liang X, Lindblom A, Lissowska J, Long J, Lu L, Magliocco AM, Martin L, McEvoy M, Meindl A, Michailidou K, Milne RL, Mints M, Montgomery GW, Nassir R, Olsson H, Orlow I, Otton G, Palles C, Perry JRB, Peto J, Pooler L, Prescott J, Proietto T, Rebbeck TR, Risch HA, Rogers PAW, Rübner M, Runnebaum I, Sacerdote C, Sarto GE, Schumacher F, Scott RJ, Setiawan VW, Shah M, Sheng X, Shu XO, Southey MC, Swerdlow AJ, Tham E, Trovik J, Turman C, Tyrer JP, Vachon C, VanDen Berg D, Vanderstichele A, Wang Z, Webb PM, Wentzensen N, Werner HMJ, Winham SJ, Wolk A, Xia L, Xiang YB, Yang HP, Yu H, Zheng W, Pharoah PDP, Dunning AM, Kraft P, De Vivo I, Tomlinson I, Easton DF, Spurdle AB, Thompson DJ. Identification of nine new susceptibility loci for endometrial cancer. Nat Commun. 2018 Aug 9;9(1):3166. doi: 10.1038/s41467-018-05427-7.

23. Belinky F, Nativ N, Stelzer G, Zimmerman S, Iny Stein T, Safran M, Lancet D. PathCards: multi-source consolidation of human biological pathways. Database (Oxford). 2015 Feb 27;2015:bav006. doi: 10.1093/database/bav006.

24. Zúñiga A, Haramis AP, McMahon AP, Zeller R. Signal relay by BMP antagonism controls the SHH/FGF4 feedback loop in vertebrate limb buds. Nature. 1999 Oct 7;401(6753):598-602. doi: 10.1038/44157.

25. Ogata H, Goto S, Fujibuchi W, Kanehisa M. Computation with the KEGG pathway database. Biosystems. 1998 Jun-Jul;47(1-2):119-28. doi: 10.1016/s0303-2647(98)00017-3.

26. Rappaport N, Twik M, Plaschkes I, Nudel R, Iny Stein T, Levitt J, Gershoni M, Morrey CP, Safran M, Lancet D. MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. Nucleic Acids Res. 2017 Jan 4;45(D1):D877-D887. doi: 10.1093/nar/gkw1012.

27. Vlachos A, Ball S, Dahl N, Alter BP, Sheth S, Ramenghi U, Meerpohl J, Karlsson S, Liu JM, Leblanc T, Paley C, Kang EM, Leder EJ, Atsidaftos E, Shimamura A, Bessler M, Glader B, Lipton JM; Participants of Sixth Annual Daniella Maria Arturi International Consensus Conference. Diagnosing and treating Diamond Blackfan anaemia: results of an international clinical consensus conference. Br J Haematol. 2008 Sep;142(6):859-76. doi: 10.1111/j.1365-2141.2008.07269.x.

28. Chevrier V, Bruel AL, Van Dam TJ, Franco B, Lo Scalzo M, Lembo F, Audebert S, Baudelet E, Isnardon D, Bole A, Borg JP, Kuentz P, Thevenon J, Burglen L, Faivre L, Rivière JB, Huynen MA, Birnbaum D, Rosnet O, Thauvin-Robinet C. OFIP/KIAA0753 forms a complex with OFD1 and FOR20 at pericentriolar satellites and centrosomes and is mutated in one individual with oral-facial-digital syndrome. Hum Mol Genet. 2016 Feb 1;25(3):497-513. doi: 10.1093/hmg/ddv488.

29. Hammarsjö A, Wang Z, Vaz R, Taylan F, Sedghi M, Girisha KM, Chitayat D, Neethukrishna K, Shannon P, Godoy R, Gowrishankar K, Lindstrand A, Nasiri J, Baktashian M, Newton PT, Guo L, Hofmeister W, Pettersson M, Chagin AS, Nishimura G, Yan L, Matsumoto N, Nordgren A, Miyake N, Grigelioniene G, Ikegawa S. Novel KIAA0753 mutations extend the phenotype of skeletal ciliopathies. Sci Rep. 2017 Nov 14;7(1):15585. doi: 10.1038/s41598-017-15442-1.

30. Inskeep KA, Zarate YA, Monteil D, Spranger J, Doherty D, Stottmann RW, Weaver KN. Genetic and phenotypic heterogeneity in KIAA0753-related ciliopathies. Am J Med Genet A. 2022 Jan;188(1):104-115. doi: 10.1002/ajmg.a.62497.